Using ultrasound articulatory signals to investigate the phonetic motivations of English /æ/ tensing

Jeff Mielke, Erik R. Thomas, and Christopher Carignan

170th ASA, Jacksonville

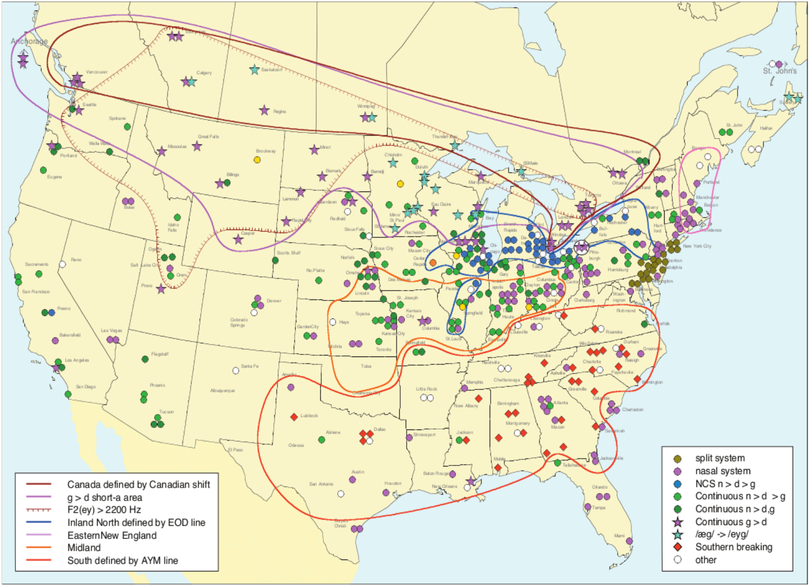
November 5, 2015

NC STATE UNIVERSITY

WESTERN SYDNEY UNIVERSITY

uOttawa

/æ/ in North American English (Labov et al., 2006)



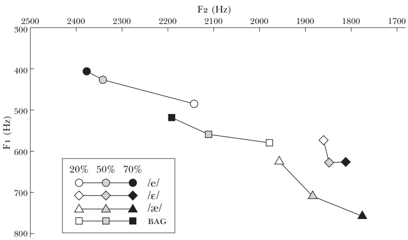
A linguistic problem: Raising of /æ/ before /g/ (1)

- Zeller (1997) reported that younger, but not older, speakers from the Milwaukee area merged /æg/ with /ejg/ (e.g., *hag*=*Haig*)
- Labov, Ash, and Boberg (2006) reported the same merger for some speakers in Wisconsin, Minnesota, and central Canada; they also noted that /æ/ tended to be higher before /g/ than before /d/ over a somewhat wider area

A linguistic problem: Raising of /æ/ before /g/ (2)

- Bauer and Parker (2008), Benson et al. (2011): speakers from Eau Claire, Wisconsin, raised /æg/
- Bauer and Parker's ultrasound data show that tongue body is raised in /æg/ but still distinct from other front vowels.
- Wassink (2015) concluded that /æg/ and /ɛg/ were raised in Seattle.

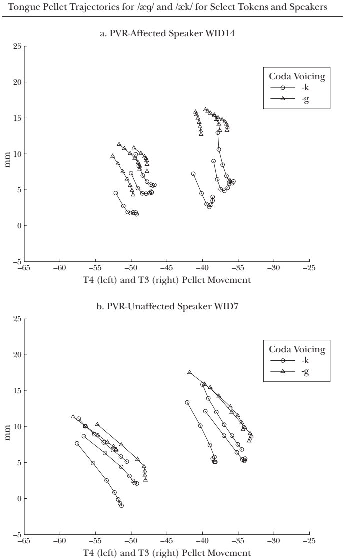
FIGURE 5
Average F1 and F2 at 20%, 50%, and 70% of Vowel Duration for /e/, /æ/, /ɛ/, and the BAG Vowel (data pooled across speakers)



A figure from Bauer and Parker (2008) illustrating differences in average trajectories

A linguistic problem: Raising of /æ/ before /g/ (3)

- Purnell (2008), using x-ray data, found that, after /æ/, Wisconsin subjects articulated /g/ more fronted than /k/ and with more forward lip position



Potential phonetic motivations for pre-velar raising

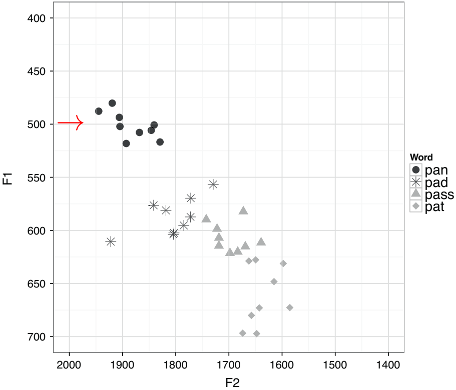
- Palatal-induced upgliding has occurred other times in the history of English, mostly before voiced stops and fricatives (and mostly not before voiceless stops).
    - Palatal [ç] conditioned upgliding in Middle English, e.g. OE *eahta* [æɒxtɑ] > *[æçtə] > ME *eight* [aiçt]
    - /g/=[ɟ], /ŋ/=[ɲ], /ʃ/, and /ʒ/, as in *bag*, *hang*, *cash*, and *azure*, respectively, condition upglides in various American dialects (see, e.g., Kurath and McDavid 1961; Hartman 1969; Thomas 2001)
- Hyperarticulation before voiceless obstruents?
    - There is some evidence that vowels can show more extreme articulations before voiceless obstruents than elsewhere (e.g., Wolf 1978; Summers 1987; Moreton 2008)
    - For low vowels, this means that F1 values are higher before voiceless obstruents than before voiced obstruents (so that the vowel reaches a lower position before voiceless obstruents).

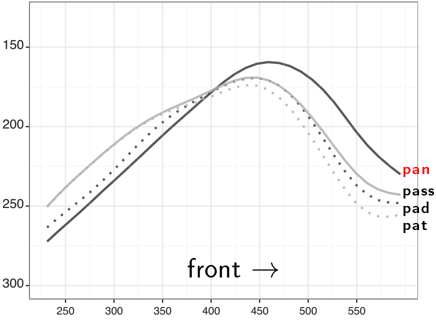
/æ/ raising in other contexts

- /æ/ raising before nasals is widespread in North American English.
- Apparent phonetic motivation: Nasalization has a strong effect on F1 in low vowels, altering their perceived height (and may also raise F2; Krakow et al. 1988)
- Raising in other contexts (e.g., before anterior voiceless fricatives) attributable to an earlier lengthening event.

De Decker and Nycz (2012): /æ/-tensing in New Jersey

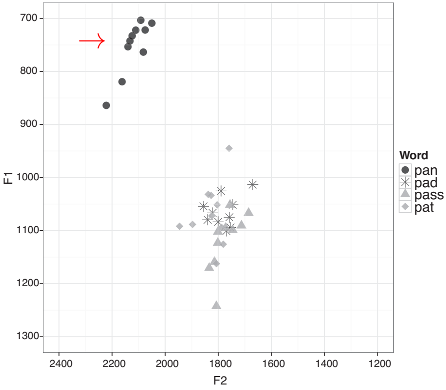
Speaker 1: acoustic raising+fronting and tongue raising+fronting

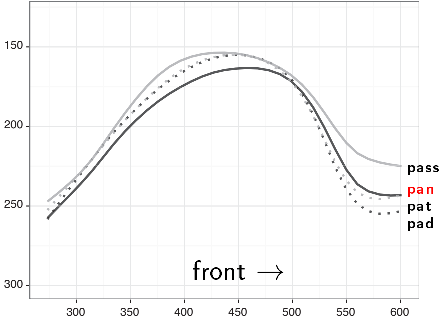




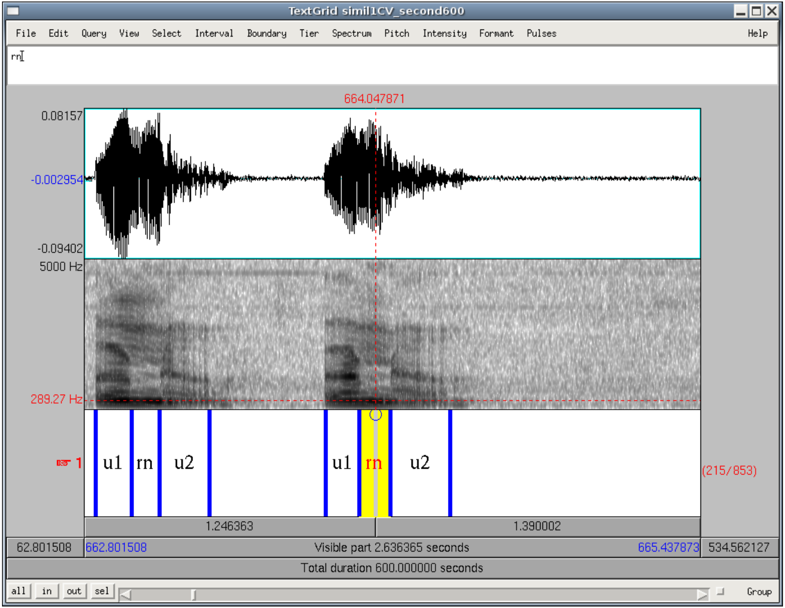
De Decker and Nycz (2012): /æ/-tensing in New Jersey

Speaker 3: acoustic raising+fronting only



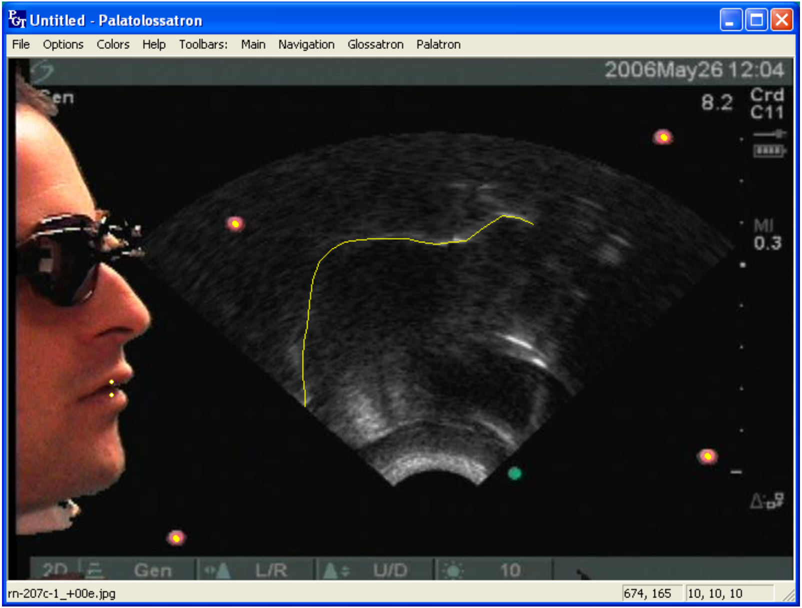


Selection of single representative image from target segment



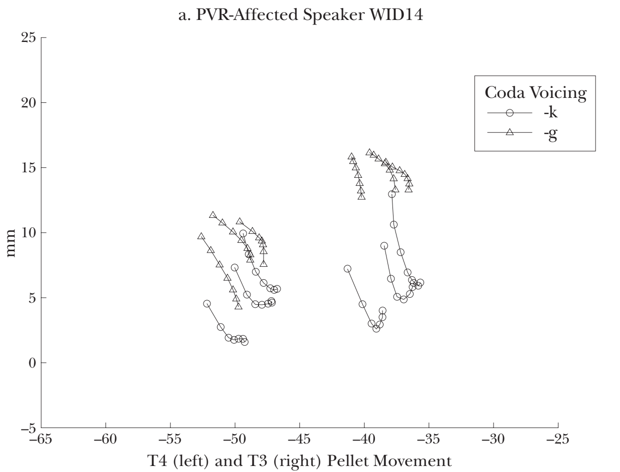
Selection of single representative image from target segment

Tongue surface contour tracing



Purnell (2008): Pellet trajectories from XRMB database (Westbury, 1994)

Tongue Pellet Trajectories for /æg/ and /æk/ for Select Tokens and Speakers



Time-varying signals from PCA of XRMB data (Story, 2007; Story and Bunton, 2013)

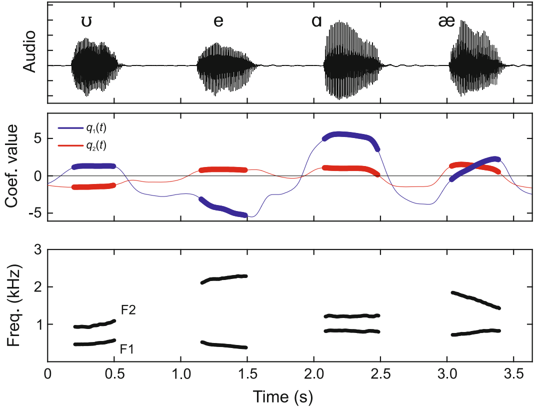
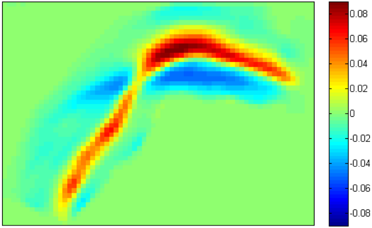


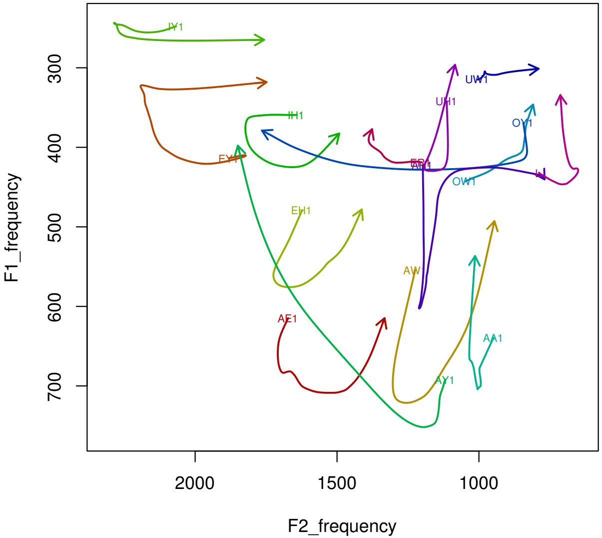
Fig. 7 Audio waveforms (*upper*), time-varying coefficients (*middle*), and formant contours (*bottom*) based on the production of four vowels by a male talker. Note the time-varying coefficients are continuous throughout the entire 3.6 s duration; the lines are thickened during the portions of time where sound is present

EigenTongues decomposition (Hueber et al., 2007)

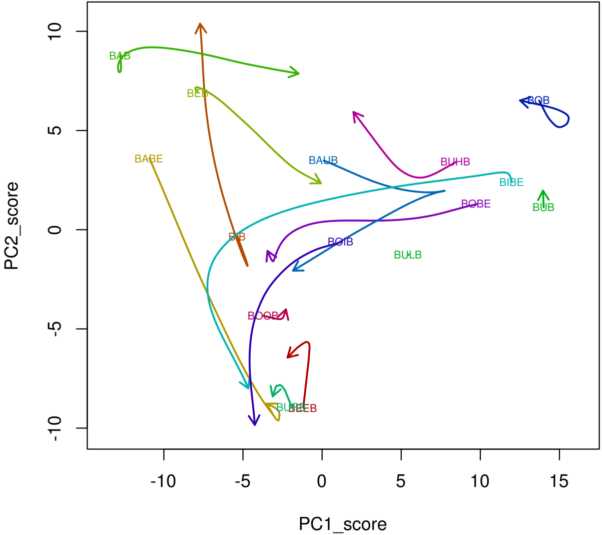


- Principal component analysis of vocal tract images (Hueber et al. 2007 for ultrasound; Carignan et al. 2013 for MRI)
- Principal Component loadings remapped onto original spatial location
- A video becomes a matrix of PC scores
- http://phon.wordpress.ncsu.edu

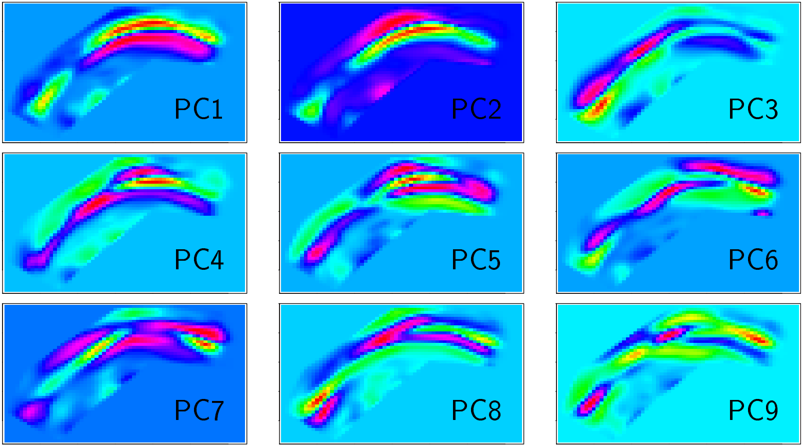
Acoustic/articulatory vowel plots



Acoustic/articulatory vowel plots

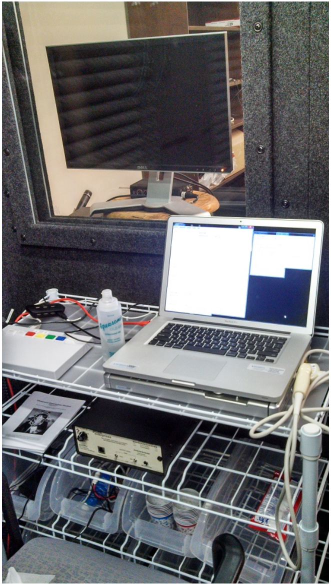


PC loadings heatmaps (first nine PCs for one speaker)



Ultrasound image acquisition (at NCSU and uOttawa)

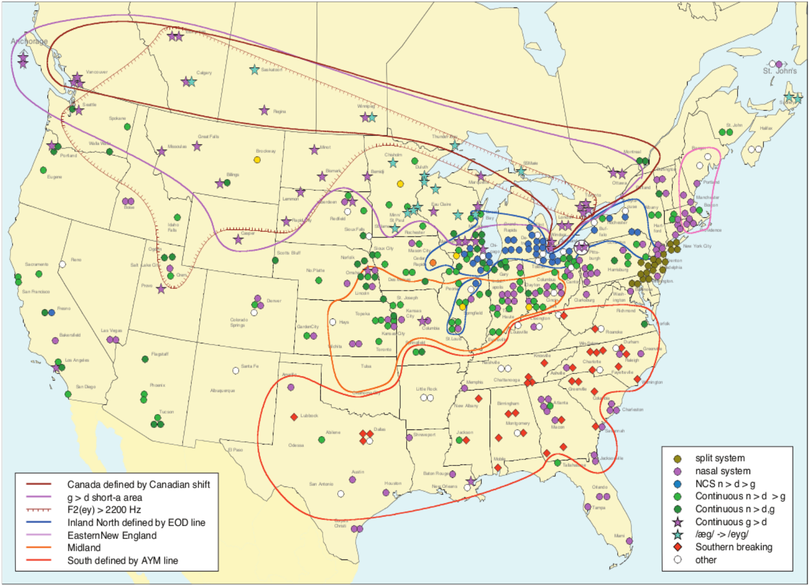
- Terason t3000
- 8MC3 microconvex array
- Ultraspeech software (Hueber et al., 2007)
- Articulate Instruments probe stabilization headset
- 120 monosyllabic words, randomized and repeated 3 times



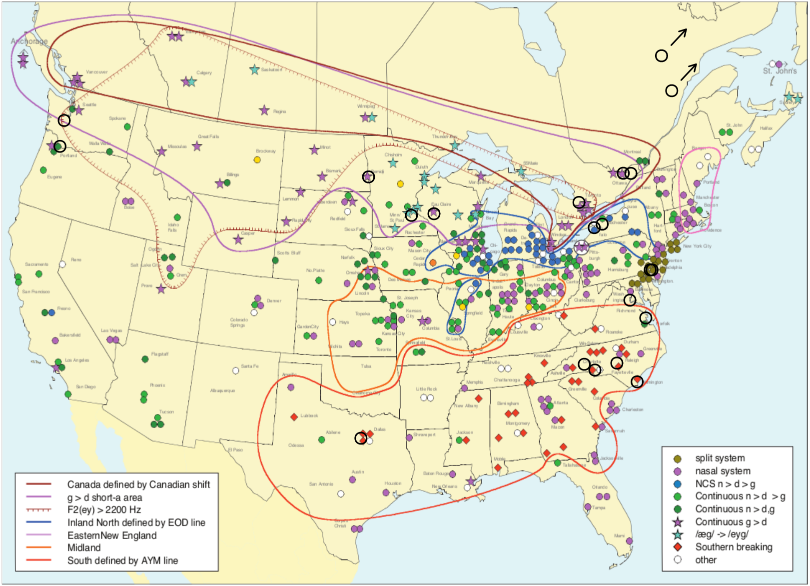
Ultrasound image acquisition (at NCSU and uOttawa)

- Terason t3000
- 8MC3 microconvex array
- Ultraspeech software (Hueber et al., 2007)
- Articulate Instruments probe stabilization headset
- 120 monosyllabic words, randomized and repeated 3 times

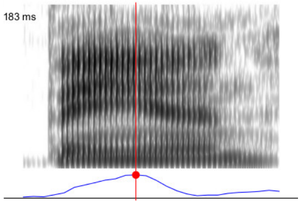
Pilot: 21 speakers (overlaid on ANAE /æ/ map; Labov et al. 2006)



Pilot: 21 speakers (overlaid on ANAE /æ/ map; Labov et al. 2006)



Time-varying articulatory signals from ultrasound

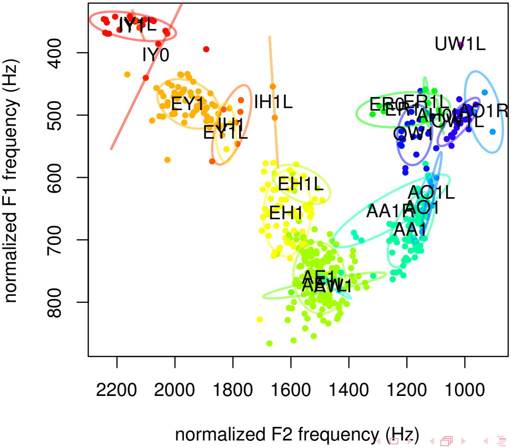


"ban"

Quantified images → articulatory signal with sampling rate = system frame rate

- Deriving time-series data from measured tongue contour tracings (Falahati, 2013)
- PCs and rotated PCs over time
- Linear Discriminant Analysis of PC scores over time (Pouplier and Hoole, 2013)
- Acoustically-inspired linear combinations of PCs over time. . .

Acoustic diagonal (Z2-Z1)

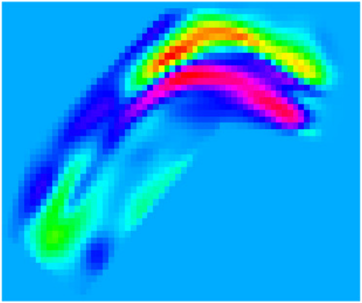


Standardized F2 − standardized F1 (Z2-Z1) = designed to match the front diagonal of the acoustic vowel space (Labov et al., 2013)

Articulatory diagonal (art.Z2Z1)

Speaker: nov03

(Broadway, NC; 1992, M)

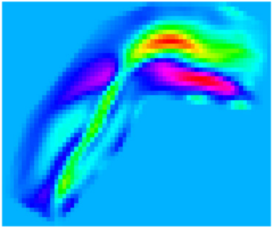


- Audio segmented using P2FA (Yuan and Liberman, 2008) and vowel/approximant formants measured at 7ms intervals
- Linear regression for each speaker's front diagonal vowels [ɑ æ ɛ ej ı i]: (Z2-Z1 ∼ PC1 + . . . + PC20)
- 20 coefficients used to make a linear weighted combination of the PCs that approximates Z2-Z1
- Second set of linear regressions using only F1 (to examine relationship between tongue position, nasalization, and F1)

Articulatory diagonal (art.Z2Z1): heatmaps

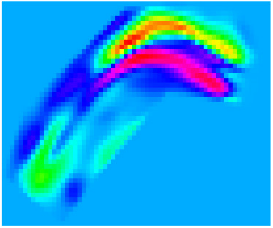
Speaker: nov01

(Vancouver, WA; 1976, M)



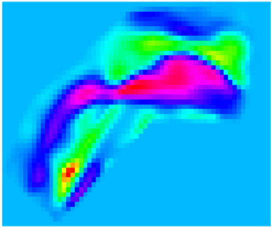
Speaker: nov03

(Broadway, NC; 1992, M)



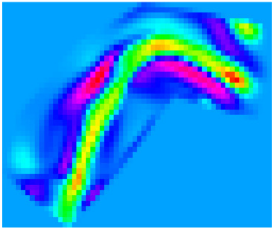
Speaker: nov04

(Olympia, WA; 1982, M)



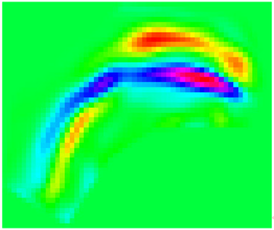
Speaker: nov07

(Arlington, TX; 1992, M)



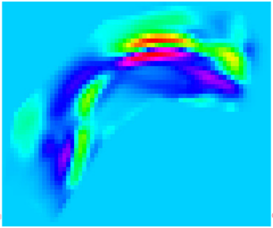
Speaker: nov11

(Wilmington, NC; 1986, M)

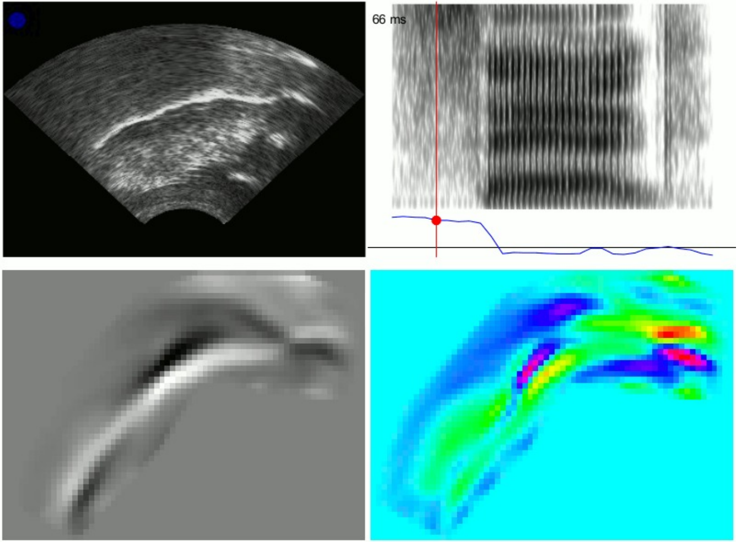


Speaker: nov12

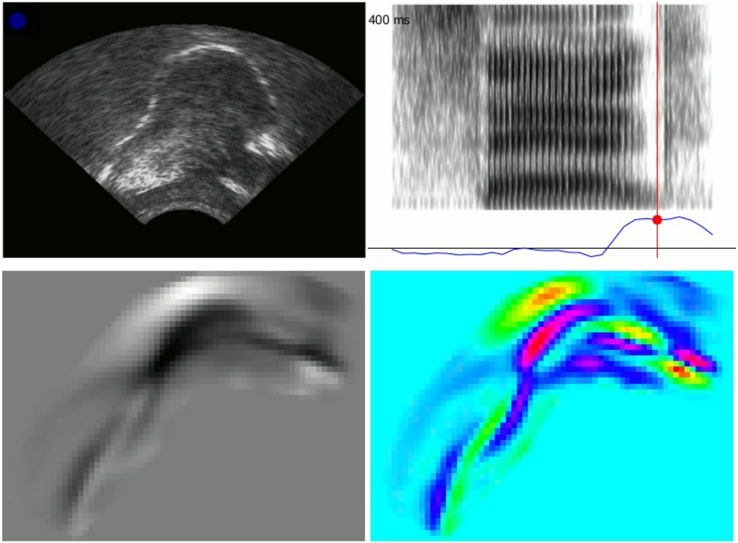
(Fargo, ND; 1981, M)



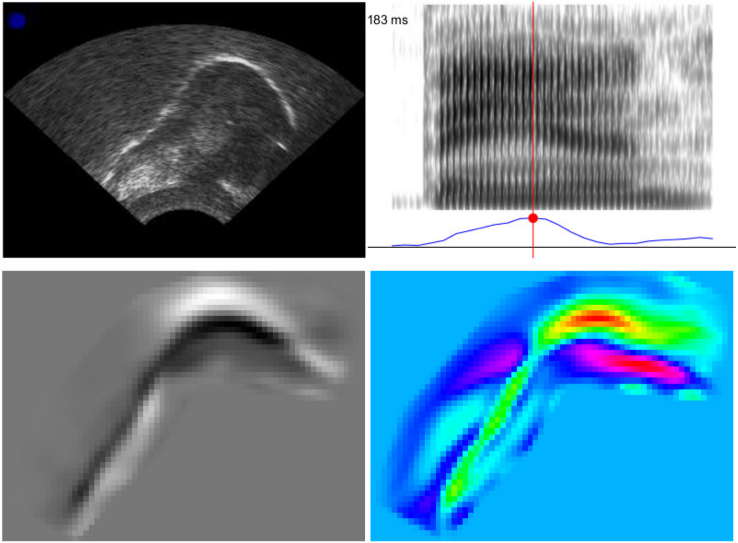
Alveolar signal (LDA with [t d n s z]): “sag”



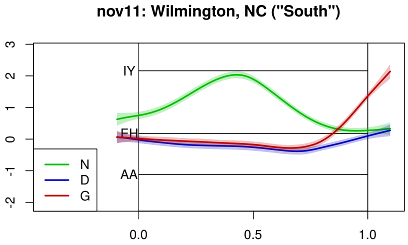
Velar signal (LDA with [k g ŋ]): "sag"



Front diagonal (art.Z2Z1) articulatory signal: “ban”

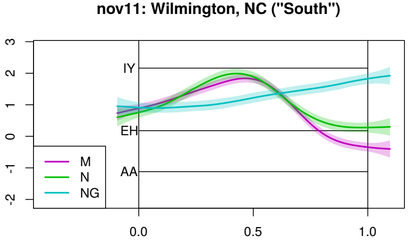


/æ/ tensing before nasals



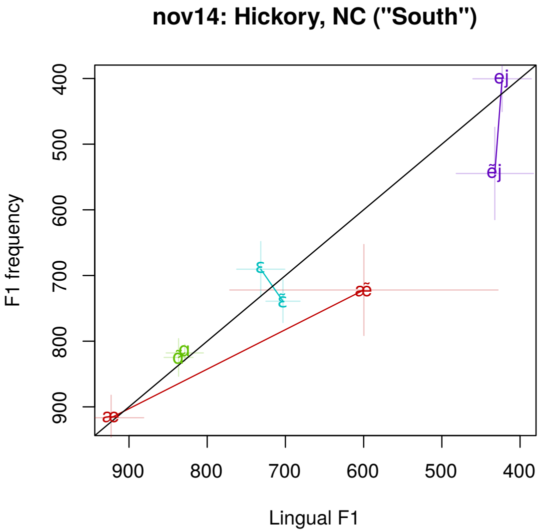
- pre-nasal tensing for all speakers except UK and Newfoundland (Wilmington, NC example)
- widespread pre-/m n/ tensing involves peak aligned ≈ with vowel nucleus
- pre-/ŋ/ tensing involves tongue raising aligned to end of vowel (anticipating following velar)
- 16/20 North Americans: pre-/ŋ/ tenser than pre-/g/, both acoustically and articulatorily (cf. Baker et al. 2008)

/æ/ tensing before nasals

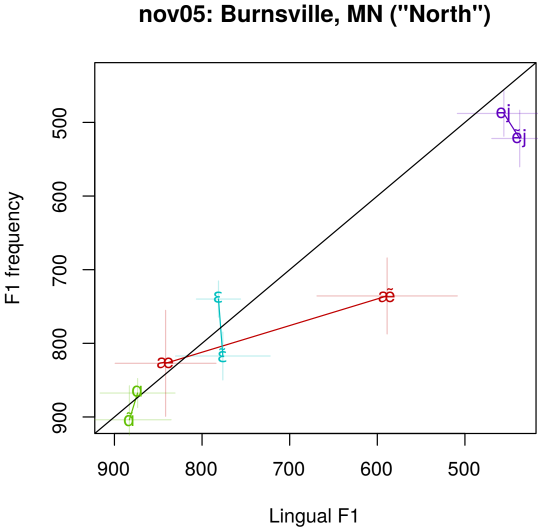


- pre-nasal tensing for all speakers except UK and Newfoundland (Wilmington, NC example)
- widespread pre-/m n/ tensing involves peak aligned ≈ with vowel nucleus
- pre-/ŋ/ tensing involves tongue raising aligned to end of vowel (anticipating following velar)
- 16/20 North Americans: pre-/ŋ/ tenser than pre-/g/, both acoustically and articulatorily (cf. Baker et al. 2008)

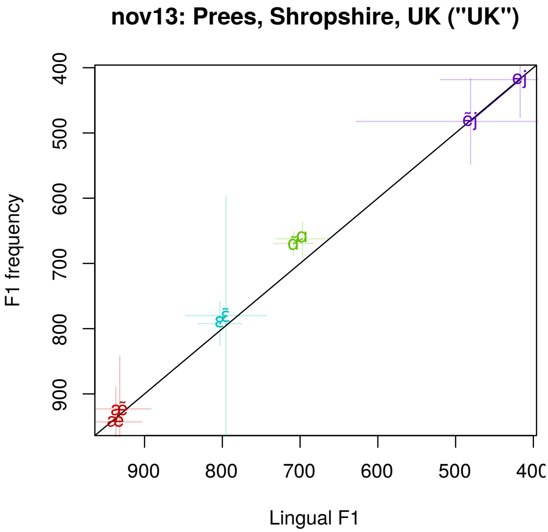
F1 vs. Lingual F1 in vowels before /m/ and /b/



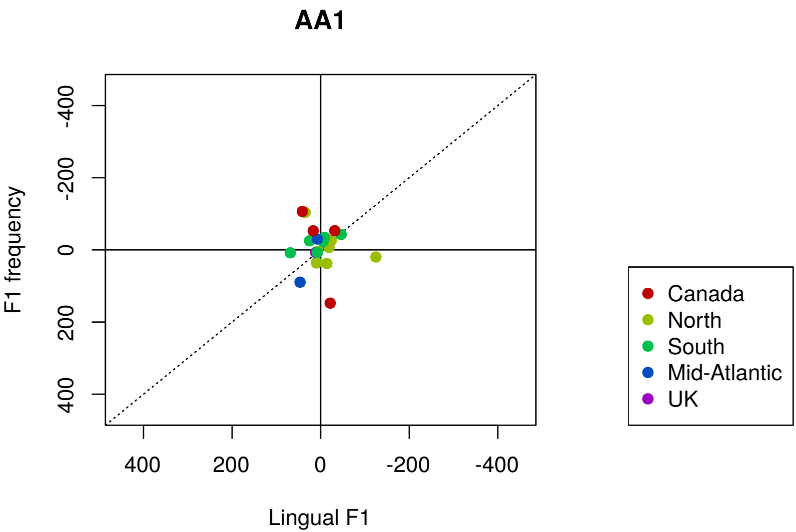
F1 vs. Lingual F1 in vowels before /m/ and /b/



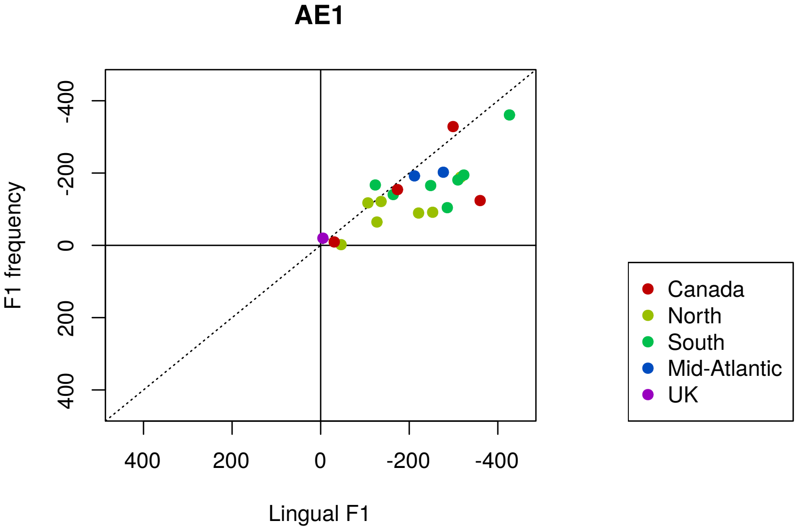
F1 vs. Lingual F1 in vowels before /m/ and /b/



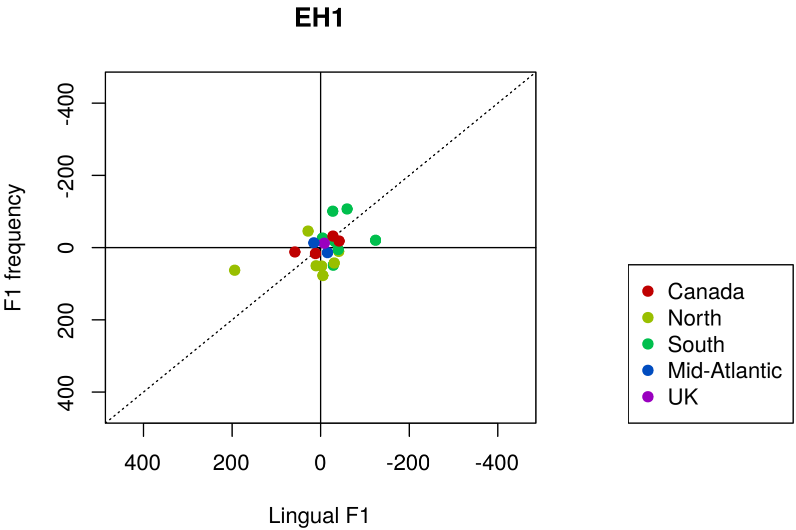
Change in F1 and Lingual F1 in pre-nasal position: /a/



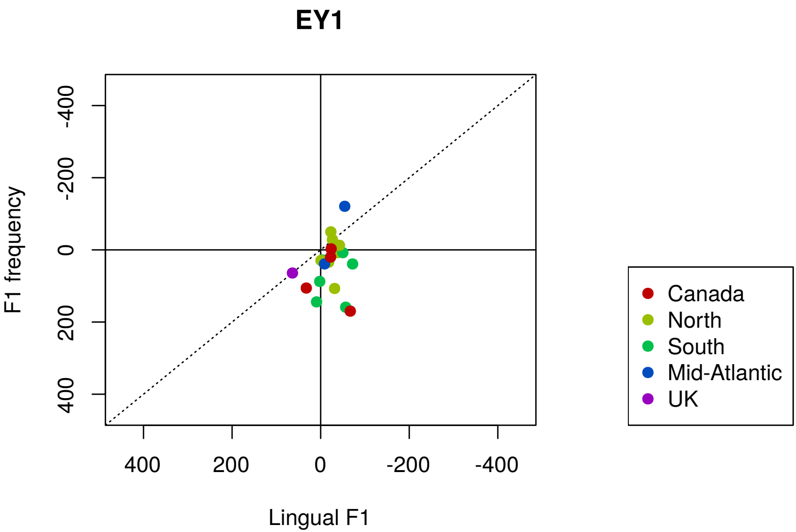
Change in F1 and Lingual F1 in pre-nasal position: /æ/



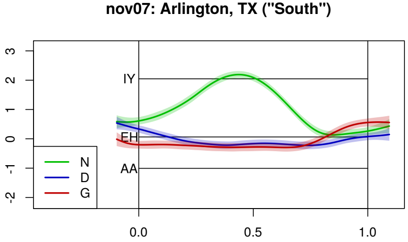
Change in F1 and Lingual F1 in pre-nasal position: /ɛ/



Change in F1 and Lingual F1 in pre-nasal position: /ej/

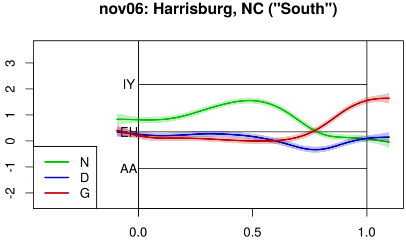


/æ/ tensing: /g/ > /d/



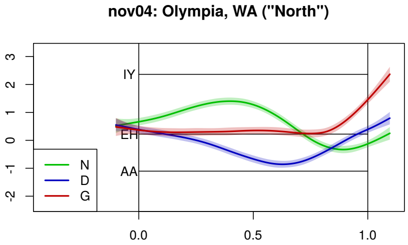
- /g/ > /d/ by end of vowel for all speakers (velar pinch) (Arlington, TX example)
- from 2nd half of vowel for most Mid-Atlantic and Southern speakers (Harrisburg, NC example)
- from 1st half of vowel for most Northern speakers (Olympia, WA example)
- entire vowel for all Ontario speakers (Barrie example)

/æ/ tensing: /g/ > /d/



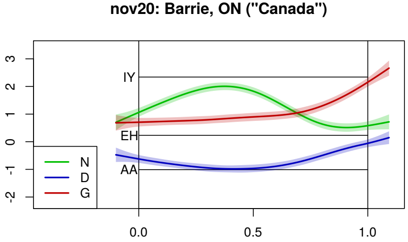
- /g/ > /d/ by end of vowel for all speakers (velar pinch) (Arlington, TX example)
- from 2nd half of vowel for most Mid-Atlantic and Southern speakers (Harrisburg, NC example)
- from 1st half of vowel for most Northern speakers (Olympia, WA example)
- entire vowel for all Ontario speakers (Barrie example)

/æ/ tensing: /g/ > /d/



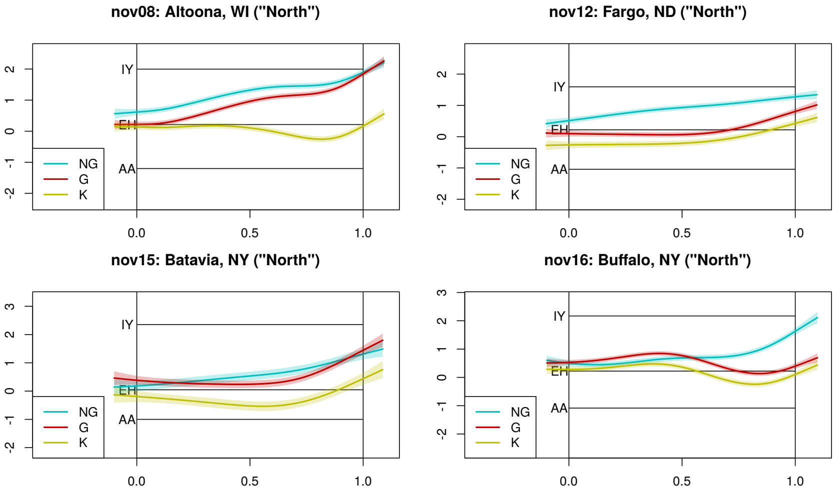
- /g/ > /d/ by end of vowel for all speakers (velar pinch) (Arlington, TX example)
- from 2nd half of vowel for most Mid-Atlantic and Southern speakers (Harrisburg, NC example)
- from 1st half of vowel for most Northern speakers (Olympia, WA example)
- entire vowel for all Ontario speakers (Barrie example)

/æ/ tensing: /g/ > /d/

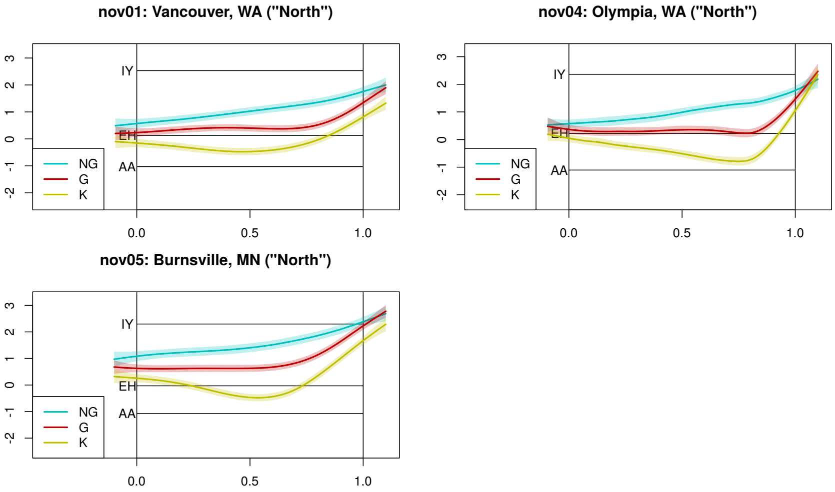


- /g/ > /d/ by end of vowel for all speakers (velar pinch) (Arlington, TX example)
- from 2nd half of vowel for most Mid-Atlantic and Southern speakers (Harrisburg, NC example)
- from 1st half of vowel for most Northern speakers (Olympia, WA example)
- entire vowel for all Ontario speakers (Barrie example)

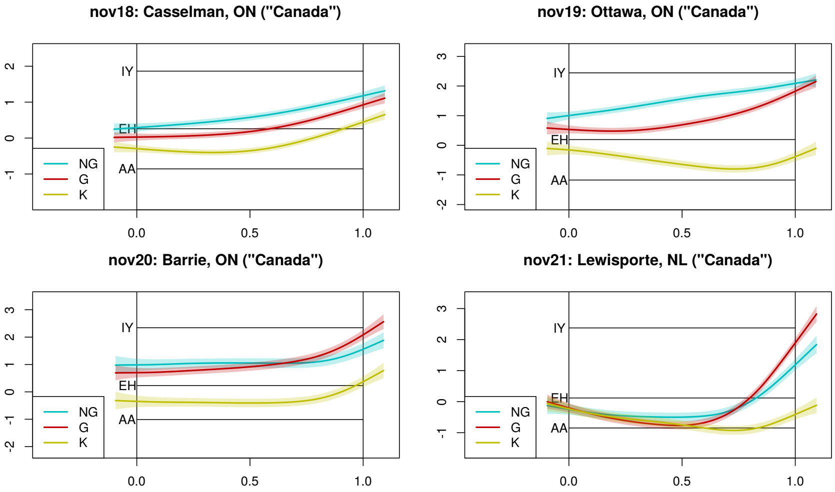
/æ/ before velars /k g ŋ/: North



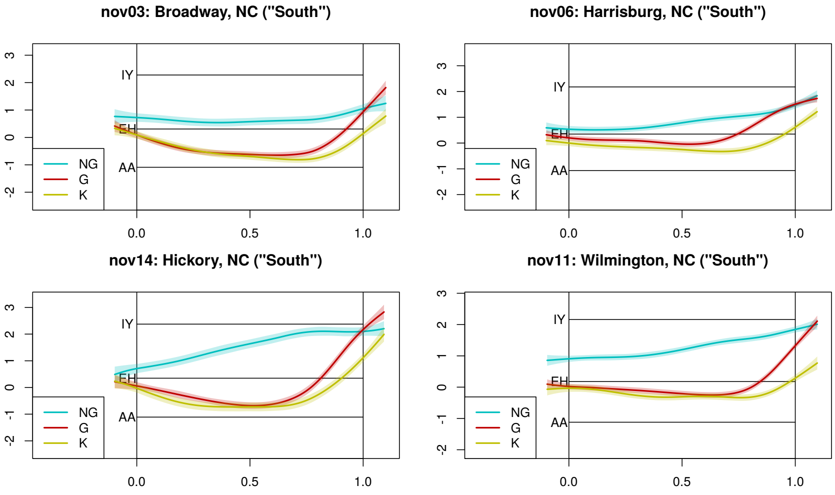
/æ/ before velars /k g ŋ/: North (including Northwest)



/æ/ before velars /k g ŋ/: Canada

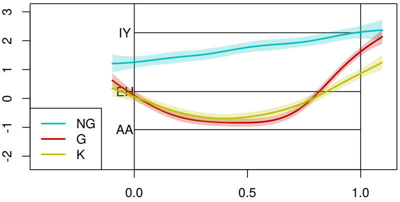


/æ/ before velars /k g ŋ/: North Carolina

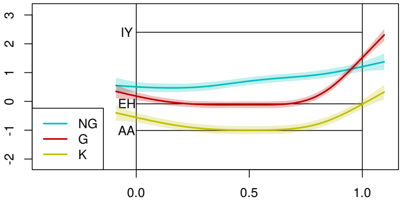


/æ/ before velars /k g ŋ/: Misc. South

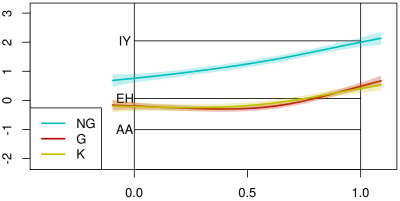
nov09: Woodbridge, VA ("South")



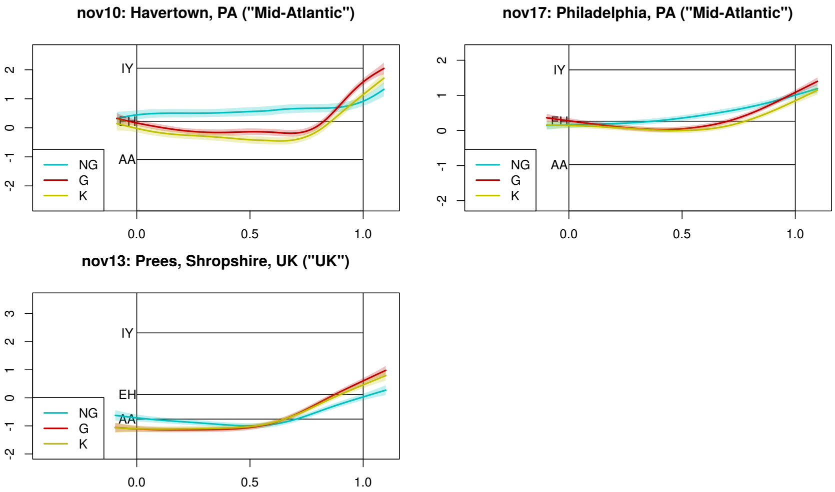
nov02: Cobbs Creek, VA ("South")



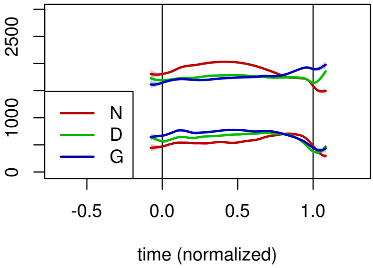
nov07: Arlington, TX ("South")

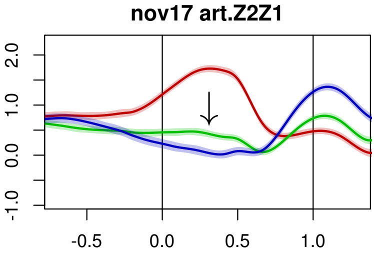


/æ/ before velars /k g ŋ/: Mid-Atlantic and UK



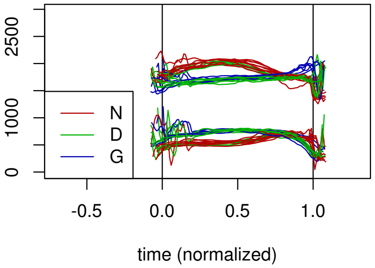
/æ/ tensing: Philadelphia

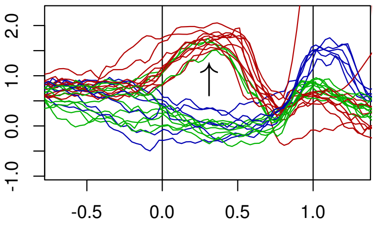




- some /d/ > /g/ for one of the Philadelphia speakers
- 'bad' > 'sad': tongue gesture similar to /æ/ before /n/ (like two of De Decker and Nycz's (2012) four New Jersey speakers)
- Anterior voiceless fricatives involve gesture similar to 'bad' and almost all of the pre-/m n/ raising we have seen so far.

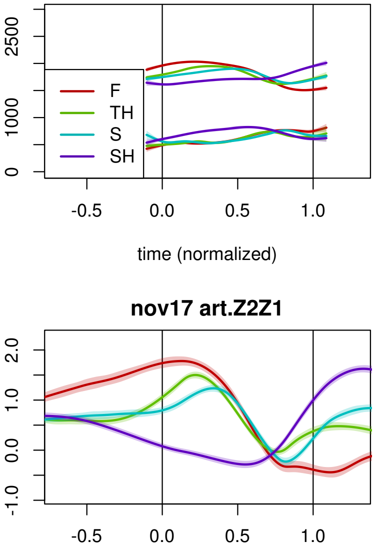
/æ/ tensing: Philadelphia





- some /d/ > /g/ for one of the Philadelphia speakers
- 'bad' > 'sad': tongue gesture similar to /æ/ before /n/ (like two of De Decker and Nycz's (2012) four New Jersey speakers)
- Anterior voiceless fricatives involve gesture similar to 'bad' and almost all of the pre-/m n/ raising we have seen so far.

/æ/ tensing: Philadelphia



- some /d/ > /g/ for one of the Philadelphia speakers
- 'bad' > 'sad': tongue gesture similar to /æ/ before /n/ (like two of De Decker and Nycz's (2012) four New Jersey speakers)
- Anterior voiceless fricatives involve gesture similar to 'bad' and almost all of the pre-/m n/ raising we have seen so far.

Summary: /æ/ raising

- Pre-nasal and Philadelphia tensing: large tongue raising gesture at the vowel nucleus
- F1 lowering in pre-nasal /æ/ is accounted for by tongue raising.
- Pre-velar /æ/ raising is a matter of timing (because pre-velar vowels end with velar contact).
- The dorsal target appears to more anterior for /g/ than for /k/ for many speakers, but conspicuously not for some, including the one UK speaker and the one Texas speaker.
    - Many of our Upper Midwest and Ontario speakers have pre-/g/ raising and articulatorily distinct /g/ and /k/.
    - Our Northwestern speakers have pre-/g/ raising but articulatorily similar /g/ and /k/.
    - Our North Carolina speakers have distinct /g/ and /k/ but no pre-/g/ raising.

Summary: articulatory signals

- Ultrasound is a *relatively* easy and practical way to collect articulatory data on the scale necessary for variation studies.
- By reducing the labor involved in ultrasound data analysis, articulatory signals make ultrasound data analysis much more flexible and make studying the dynamics of speech production more practical.
- Signals derived from PCs using acoustic data can be used to track linguistically relevant tongue movements (e.g. articulatory movement along the front edge of the vowel space).
- Acoustically-derived signals can also be used to distinguish effects of tongue movement from effects of lips, nasalization, etc.

Thanks

- Data collection and analysis at North Carolina State University was made possible by funding from the NCSU Department of English and the College of Humanities and Social Sciences.
- Data collection at the University of Ottawa was made possible by CFI grant #15834 "Sound Patterns Laboratory/Laboratoire des structures sonores" to Jeff Mielke and Marc Brunelle.
- Analysis has been supported by NSF grant #BCS 1451475 "Phonological implications of covert articulatory variation".
- Thanks to Robin Dodsworth and Elliott Moreton for discussion.

NC STATE UNIVERSITY NSF uOttawa

Articulatory signals Matlab scripts and polar SSANOVA R script: http://phon.wordpress.ncsu.edu or google "NCSU phonology"

Baker, Adam, Jeff Mielke, and Diana Archangeli. 2008. More velar than /g/: Consonant coarticulation as a cause of diphthongization. Somerville, MA: Cascadilla Proceedings Project. URL http://www.lingref.com/cpp/wccfl/26/paper1656.pdf.

Bauer, Matt, and Frank Parker. 2008. /æ/-raising in Wisconsin English. *American Speech* 83:403–431.

Benson, Erica J., Michael J. Fox, and Jared Balkman. 2011. The bag that scott bought: The low vowels in northwest Wisconsin. *American Speech* 86:271–311.

Carignan, C., R. Shosted, M. Fu, Z.-P. Liang, and B. Sutton. 2013. The role of the tongue and pharynx in enhancement of vowel nasalization: A real-time MRI investigation of French nasal vowels. In *Proceedings of INTERSPEECH 2013*, 3042–3046.

De Decker, Paul M., and Jennifer R. Nycz. 2012. Are tense [æ]s really tense? the mapping between articulation and acoustics. *Lingua* 122:810–821.

Falahati, Reza. 2013. Gradient and categorical consonant cluster simplification in Persian: An ultrasound and acoustic study. Doctoral Dissertation, University of Ottawa.

Hartman, James W. 1969. Some preliminary findings from DARE. *American Speech* 44:191–199.

Hueber, T., G. Aversano, G. Chollet, B. Denby, G. Dreyfus, Y. Oussar, P. Roussel, and M. Stone. 2007. Eigentongue feature extraction for an ultrasound-based silent speech interface. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, 1245–1248. Honolulu, HI: Cascadilla Press.

Krakow, R.A., P.S. Beddor, L. Goldstein, and C.A. Fowler. 1988. Coarticulatory influences on the perceived height of nasal vowels. *Journal of the Acoustical Society of America* 83:1146–1158.

Kurath, Hans, and Raven I. Jr. McDavid. 1961. *The pronunciation of English in the Atlantic States*. Ann Arbor: University of Michigan Press. Publication of the American Dialect Society 85.

Labov, William, Sharon Ash, and Charles Boberg. 2006. *The atlas of North American English: Phonetics, phonology and sound change*. De Gruyter Mouton.

Labov, William, Ingrid Rosenfelder, and Josef Fruehwald. 2013. One hundred years of sound change in Philadelphia: Linear incrementation, reversal, and reanalysis. *Language* 89:30–65.

Moreton, Elliott. 2008. Realization of the english postvocalic [voice] contrast in f1 and f2. *Journal of Phonetics* 32:1–33.

Pouplier, Marianne, and Phillip Hoole. 2013. Comparing principal component analysis of ultrasound images with contour analyses in a study of tongue body control during german coronals. Ultrafest VI, Edinburgh, November 2013.

Purnell, Thomas C. 2008. Prevelar raising and phonetic conditioning: Role of labial and anterior tongue gestures. *American Speech* 83:373–402.

Story, Brad H. 2007. Time dependence of vocal tract modes during production of vowels and vowel sequences. *Journal of the Acoustical Society of America* 121:3770–3789.

Story, Brad H., and Kate Bunton. 2013. Simulation and identification of vowels based on a time-varying model of the vocal tract area function. In *Vowel inherent spectral change*, ed. G. S. Morrison and P. F. Assmann, 155–174. Berlin: Springer-Verlag.

Summers, W. Van. 1987. Effects of stress and final-consonant voicing on vowel production: articulatory and acoustic analyses. *Journal of the Acoustical Society of America* 82:847–863.

Thomas, Erik R. 2001. *An acoustic analysis of vowel variation in New World English*. Durham, N.C.: Duke University Press. Publication of the American Dialect Society 85.

Wassink, Alicia Beckford. 2015. Sociolinguistic patterns in seattle english. *Language Variation and Change* 27:31–58.

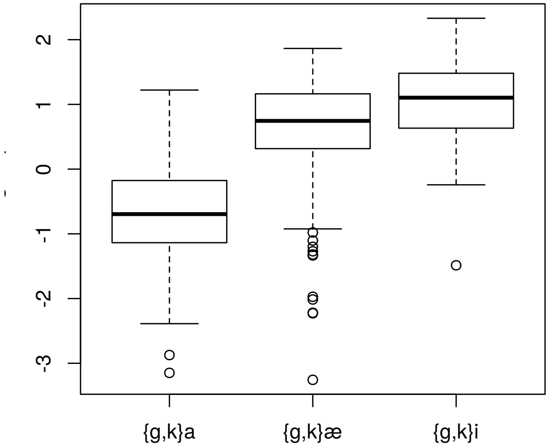
Westbury, John R. 1994. X-ray microbeam speech production database user's handbook. University of Wisconsin, Madison, WI.

Wolf, Catherine G. 1978. Voicing cues in english final stops. *Journal of Phonetics* 6:299–309.

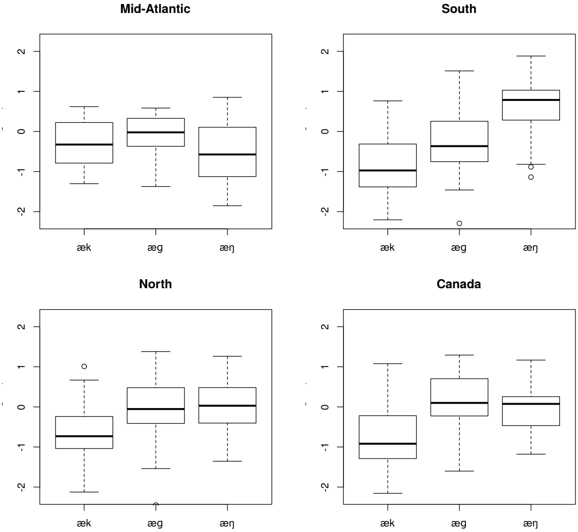
Yuan, Jiahong, and Mark Liberman. 2008. Speaker identification on the SCOTUS corpus. In *Proceedings of Acoustics '08*, 5687–5690.

Zeller, Christine. 1997. The investigation of a sound change in progress: /æ/ to /e/ in Midwestern American English. *Journal of English Linguistics* 25:142–155.

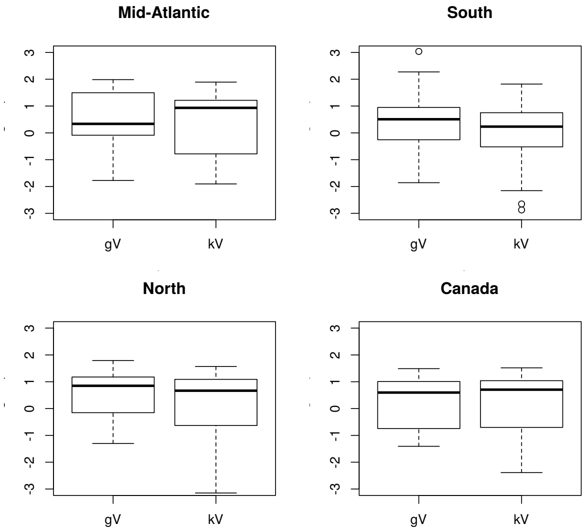
Anteriority of closure: Velar palatalization



Anteriority of closure: Velars after /æ/



Anteriority of closure: Velars before /i ej ɛ ɑ/



Anteriority of closure: Velars after /i ej ɛ ɑ/